REMARKS

Claims 1, 2, 4-11, 13-14, 16-19 are pending. Claims 1, 10 and 13 have been amended to overcome the non-statutory subject matter rejection, and not in view of the prior art. As a result of the amendments, claims 3, 12 and 15 have been canceled as being duplicative, and not in view of the prior art. Figures 2 and 3 have been amended by designating each of these two figures with the legend – Prior Art. For the reasons discussed in detail below, all of the pending claims are in condition for allowance.

Prior Art Rejections

The Office action has rejected claims 1-19 under 35 U.S.C. § 103(a) over "Reconstruction And Prediction of Nonlinear Dynamical Systems: A Hierarchical Bayes Approach With Neural Nets" to Matsumoto et al. ("Matsumoto") in view of U.S. Patent 5,692,107 to Simoudis et al. ("Simoudis"). Applicants respectfully traverse these rejections. In the following, applicants provide an overview or their invention and then discuss the differences with the relied-upon references.

Applicants' technique for modeling a data set uses a variation of a probabilistic model known as a Relevance Vector Machine (RVM). This variation uses product approximations, including the distribution of hyperparameters of the RVM, to obtain a posterior distribution. The RVM uses a prior distribution for the data set to infer the resulting prediction model. This prior distribution is determined from selection of an initial set of hyperparameters. In this version of

applicants' technique, the RVM generates a separate distribution for each

hyperparameter and iteratively updates the distribution of the set of

hyperparameters, the distribution of the set of weights, and the distribution of the

set of predetermined additional parameters. The iteration ceases upon reaching

a chosen convergence criterion. Applicants' technique results in constructing a

model that outputs a posterior distribution over both parameters and

hyperparameters that is used for probabilistic prediction of an event(s) or

expected behavior(s) for a given input(s). Note that the above description is for

example and informational purposes only, and should not be used to interpret the

claims, which are discussed below.

Applicants respectfully disagree that claims 1-19 are unpatentable over

Matsumoto in view of Simoudis. First of all, Matsumoto is directed toward a

hierarchical bayesian approach to nonlinear time series prediction weighted by

marginal likelihoods. Although Matsumoto may describe using a family of prior

distributions parameterized by hyperparameters in the nonlinear time series

prediction, Matsumoto, however, teaches computing three discrete levels of

posterior distributions in order to compute a predictive distribution. (Column 2,

page 1061 and column 4 page 1062.) Matzumoto describes calculating the

posterior distribution of the weight parameters first to obtain the most probable

weight parameters, and then using the most probable weight parameters to

calculate the posterior distribution of the hyperparameters next to obtain the most

probable hyperparameters. Then, Matzumoto may calculate the posterior

distribution for the most probable model using the most probable weight

parameters and the most probable hyperparameters. Finally, the predictive

distribution may be calculated using the most probable model, the most probable

hyperparameters and the most probable weight parameters. Significantly,

Matzumoto thus teaches away from updating both the set of weights and the set

of hyperparameters iteratively to approximate the posterior distribution for the

model.

As acknowledged by the Office Action, Matsumoto fails to disclose the

limitations of "interactively updating the distribution of the set of weights, the

distribution of the set of hyperparameters, and the distribution of the set of

predetermined additional parameters until a predetermined convergence criterion

has been reached, such that the product of the distribution of the set of

hyperparameters, the distribution of the set of weights, and the distribution of the

set of predetermined additional parameters as have been iteratively updated

approximates the posterior distribution for modeling of the data set for

probabilistic prediction," as recited in claims 1 and 13.

The Office action surprisingly contends that Simoudis discloses these

limitations as recited in claims 1 and 13. Simoudis, however, is directed to

extracting data to generate predictive models. Simoudis generally teaches the

integration of top-down and bottom-up data mining to generate predictive

models. An exemplary analysis module for top-down or bottom-up data mining

includes custom designed or commercially available deductive database

processing, inductive learning, clustering, case-based reasoning, visualization,

and statistical analysis. Such modules may be added or omitted. (Simoudis,

column 3, lines 21-30.) A user may select a data mining module and module

specific data files are generated and stored (Simoudis, column 4, lines 43-45).

A predictive model may be extracted and saved if the results are satisfactory to a

user. (Simoudis, column 4, lines 51-58.)

Significantly, Simoudis does not relate to a variation approach to a

relevance vector learning machine. Instead, Simoudis describes using

commercially available or customized top-down and bottom-up data mining

analysis to generate predictive models. Column 4, lines 42-57, of Simoudis do

not disclose "interactively updating the distribution of the set of weights, the

distribution of the set of hyperparameters, and the distribution of the set of

predetermined additional parameters until a predetermined convergence criterion

has been reached, such that the product of the distribution of the set of

hyperparameters, the distribution of the set of weights, and the distribution of the

set of predetermined additional parameters as have been iteratively updated

approximates the posterior distribution for modeling of the data set for

probabilistic prediction," as alleged by the Office action. This section instead

discloses how a user may select a data analysis module to perform data mining.

In specific, this section discloses that module-specific data files and a data

specification are stored upon selecting the data mining module. A GUI for the

selected module is then used to set module-specific parameters, which may be in

the form of user queries. The selected data analysis module then is applied to

the target data set and the results are returned to the user for examination via the

module GUI. A predictive model may be extracted if the results are satisfactory to the user.

Nowhere in this section or elsewhere in Simoudis can there be found any disclosure of interactively updating the distribution of the set of weights, the distribution of the set of hyperparameters, and the distribution of the set of predetermined additional parameters until a predetermined convergence criterion has been reached, as recited in claims 1, 13, 10, 16.

Without any disclosure of such claimed limitations, Simoudis could not possibly suggest, as the Office action alleges, that Matsumoto could modify the hierarchical bayesian approach to nonlinear time series prediction to include "interactively updating the distribution of the set of weights, the distribution of the set of hyperparameters, and the distribution of the set of predetermined additional parameters until a predetermined convergence criterion has been reached." Significantly, there is no such discussion to be found in Simoudis. Simoudis does not even discuss any distribution of a set of weights, much less a distribution of a set of hyperparameters, or any set of predetermined additional parameters.

For at least these reasons, applicants submit that claims 1, 13, 10, and 16 are patentable over the prior art of record. Applicants also submit that their dependent claims 2, 4-5, 11, 14, and 17-19 by similar analysis are patentable over the prior art of record.

As acknowledge by the Office Action, Matsumoto also fails to disclose the limitations of "determining a relevance vector learning machine via a variational

approach to obtain a posterior distribution for the data set," as recited in claims 6

and 16. Equally surprisingly, the Office action contends that the inductive

module 104 of Fig. 1 in Simoudis discloses these limitations. The inductive

module 104' of Fig. 1 in Simoudis is a custom designed or commercially available

bottom-up data analysis module. (Simoudis, column 2, lines 40 -42 and column

3, lines 20-25.) Nowhere in the alleged section or elsewhere in Simoudis is there

any disclosure of "determining a relevance vector learning machine via a

variational approach to obtain a posterior distribution for the data set," as recited

in claims 6 and 16. Without any disclosure of such claimed limitations, Simoudis

could not possibly suggest, as the Office action alleges, that Matsumoto could

modify the hierarchical bayesian approach to nonlinear time series prediction to

include "determining a relevance vector learning machine via a variational

approach to obtain a posterior distribution for the data set." Simoudis does not

discuss determining a relevance vector learning machine, much less determining

a relevance vector learning machine via a variational approach.

For at least these reasons, applicants submit that claims 6 and 16, and

their dependent claims 7-9 and 17-19 by similar analysis, are patentable over the

prior art of record.

As acknowledge by the Office Action, Matsumoto further fails to disclose

the limitation of "interactively updating the distribution of the set of weights, the

distribution of the set of hyperparameters, and the distribution of the set of

parameters accounting for noise until a predetermined convergence criterion has

been reached, such that the product of the distribution of the set of

hyperparameters, the distribution of the set of weights, and the distribution of the

set of parameters accounting for noise as have been iteratively updated

approximates the posterior distribution for modeling of the continuous data set for

probabilistic prediction," as recited in claim 10. Once again, surprisingly, the

Office action contends that column 4, lines 42-57 of Simoudis specifically

disclose these limitations. Column 4, lines 42-57, of Simoudis do not specifically

disclose "interactively updating the distribution of the set of weights, the

distribution of the set of hyperparameters, and the distribution of the set of

parameters accounting for noise until a predetermined convergence criterion has

been reached, such that the product of the distribution of the set of

hyperparameters, the distribution of the set of weights, and the distribution of the

set of parameters accounting for noise as have been iteratively updated

approximates the posterior distribution for modeling of the continuous data set for

probabilistic prediction," as alleged by the Office action. This section instead

discloses how a user may select a data analysis module to perform data mining.

In specific, this section discloses that module-specific data files and a data

specification are stored upon selecting the data mining module. A GUI for the

selected module is then used to set module-specific parameters, which may be in

the form of user queries. The selected data analysis module then is applied to

the target data set and the results are returned to the user for examination via the

module GUI. A predictive model may be extracted if the results are satisfactory

to the user.

Nowhere in this section or elsewhere in Simoudis can there be found any disclosure of interactively updating the distribution of the set of weights, the distribution of the set of hyperparameters, and the distribution of the set of parameters accounting for noise until a predetermined convergence criterion has been reached, as recited in claim 10. Without any disclosure of such claimed limitations, Simoudis could not possibly suggest, as the Office action alleges, that Matsumoto could modify the hierarchical bayesian approach to nonlinear time series prediction to include "interactively updating the distribution of the set of weights, the distribution of the set of hyperparameters, and the distribution of the set of parameters accounting for noise until a predetermined convergence criterion has been reached." Significantly, Simoudis does not even discuss any distribution of a set of weights, much less a distribution of a set of hyperparameters, or any set of parameters accounting for noise.

For at least these reasons, applicants submit that claim 10, and dependent claim 11 by similar analysis, are patentable over the prior art of record.

To establish *prima facie* obviousness of a claimed invention, all of the claim limitations must be taught or suggested by the prior art; (*In re Royka*, 490 F.2d 981, 180 USPQ 580 (CCPA 1974)), and "all words in a claim must be considered in judging the patentability of that claim against the prior art;" (*In re Wilson*, 424 F.2d 1382, 1385, 165 USPQ 494, 496 (CCPA 1970)). Neither Matusmoto nor Simoudis, taken alone or together, teach or suggest all of the claim limitations of applicants' invention. The Office action acknowledges that

Matsumoto does not disclose all of the claim limitations. Moreoever, nowhere in Simoudis can there be found any discussion of claim limitations that the Office action has already acknowledged are not disclosed by Matsumoto. Without any disclosure of such claim limitation in either Matsumoto or Simoudis, all of the claim limitations could not possibly be either taught or suggested by combining the references. For at least the foregoing reasons, factual and legal, applicants submit that neither Matsumoto nor Simoudis, whether considered alone or in any permissible combination, meet these requirements, and thus that the present Office action has failed to establish prima facie obviousness as a matter of law with respect to any of the claimed subject matter.

Even if Simoudis had disclosed the missing claim limitations, there is no motivation or suggestion in Simoudis for modifying Matsumoto. Contending that the missing claim limitations exist in Simoudis and that Simoudis provides the motivation for modifying Masumoto to produce applicants' invention could only have been made by applying applicants' teaching. Combining prior art references without evidence of such a suggestion, teaching, or motivation simply takes the inventor's disclosure as a blueprint for piecing together the prior art to defeat patentability--the essence of impermissible hindsight. *In re Dembiczak*, 175 F.3d 994, 999, 50 USPQ2d 1614, 1617 (Fed. Cir. 1999). Furthermore, Matsumoto teaches computing three discrete levels of posterior distributions in order to compute a predictive distribution. Thus, Matzumoto teaches away from updating both the set of weights and the set of hyperparameters iteratively to approximate the posterior distribution for the model.

Reconsideration and withdrawal of the rejections of pending claims based on Matsumoto and/or Simoudis is respectfully requested.

Non-statutory Subject Matter Rejections

The Office action has also rejected claims 1, 10, and 13 under 35 U.S.C. § 101 as being directed to non-statutory subject matter. Applicants strongly disagree with this rejection. Applicants' technique constructs a probabilistic model that takes real-world data and outputs a posterior distribution used for probabilistic prediction of an event or expected behavior. Claim 1 recites the limitation "for modeling of the data set for probabilistic prediction." Claim 1, therefore, explicitly recites limitations to a practical application. This is but one example of how the claims are statutory. However, to expedite allowance of this patent application, applicants have amended claims 1, 10, and 13 to recite the additional limitation of "outputting at least the product approximating the posterior distribution for modeling of the data set" or similar language. These claims as amended now explicitly recite additional limitations to a practical application, and the method, as now claimed, produces a concrete, tangible and useful result. See *AT&T Corp. v. Excel Communications, Inc.*, 172 F. 3d 1352, 1358, 50 USPQ2d 1447, 1452 (Fed. Cir. 1999).

Thus, the claimed invention as a whole clearly accomplishes a practical application, as it produces a useful, concrete and tangible result, and does so without pre-empting other uses of the mathematical principle behind it. *State*

Street Bank & Trust Co. v. Signature Financial Group Inc., 149 F. 3d 1368, 1374,

47 USPQ2d 1596, 1601-02 (Fed. Cir. 1998); *AT&T Corp. v. Excel*

Communications, Inc., 172 F. 3d 1352, 1358, 50 USPQ2d 1447, 1452 (Fed. Cir.

1999). Reconsideration and withdrawal of the 35 U.S.C. § 101 rejections is

respectfully requested.

In re application of TIPPING ET AL.
Serial No. 09/520,678

Conclusion

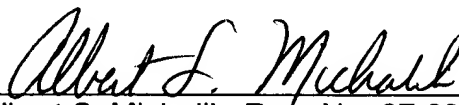In view of the foregoing remarks, it is respectfully submitted that claims 1-19

of the present application are patentable over the prior art of record, and that the

application is in good and proper form for allowance. A favorable action on the part

of the Examiner is earnestly solicited.

If in the opinion of the Examiner a telephone conference would expedite the

prosecution of the subject application, the Examiner is invited to call the

undersigned attorney at (425) 836-3030.

Respectfully submitted,

Albert S. Michalik, Reg. No. 37,395
Attorney for Applicants
Law Offices of Albert S. Michalik, PLLC
704 - 228th Avenue NE, Suite 193
Sammamish, WA 98074
(425) 836-3030
(425) 836-8957 (facsimile)

21